BIOTE( )LOGY
SYSTEMS
BRANCH

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: *09/757,333*

Source: *OIPE*

Date Processed by STIC: *1/25/2001*

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE**
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A**
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER**
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program
employing a logical and intuitive user-interface to check whether a sequence listing is in
compliance with format and content rules. Checker Version 3.0 works for sequence listings
generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old
rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual
Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-
compliant. Checker allows public users to check sequence listings in Computer Readable form
(CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence
listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/757,333

DATE: 01/25/2001
TIME: 11:25:05

Input Set : **A:\sequence listing.txt**
Output Set: **N:\CRF3\01252001\I757333.raw**

```
     3 <110> APPLICANT: APPLICANT: Samuel I. Achilefu
     4       Raghavan Rajagopalan
     5       Richard B. Dorshow
     6       Joseph E. Bugaj
     7       ASSIGNEE: Mallinckrodt Inc.
     9 <120> TITLE OF INVENTION: TITLE: Versatile Hydrophilic Dyes
    11 <130> FILE REFERENCE: DOCKET/FILE REFERENCE: MRD-67
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/757,333
C--> 13 <141> CURRENT FILING DATE: 2001-01-09
    13 <150> PRIOR APPLICATION NUMBER: PRIOR APPLICATION NUMBER: 09/484,321
W--> 14 <151> PRIOR FILING DATE: FILING DATE: 2000-01-18
E--> 16 <160> NUMBER OF SEQ ID NOS: NUMBER OF SEQUENCES: 8
    18 <170> SOFTWARE: SOFTWARE: FastSEQ for Windows Version 3.0
```

Does Not Comply
Corrected Diskette Needed

ERRORED SEQUENCES

```
E--> 20 <210> SEQ ID NO: SEQ ID NO:1
    21 <211> LENGTH: LENGTH: 8
E--> 22 <212> TYPE: TYPE: PRT
    23 <213> ORGANISM: ORGANISM:Synthetic
    25 <221> NAME/KEY: MOD_RES
    26 <222> LOCATION: (1)...(0)
    27 <223> OTHER INFORMATION: Xaa = D-Phe
E--> 28 <224> Xbb = Cys with an intramolecular disulfide bond
E--> 30 <225> Xcc = D-Trp
E--> 32 <400> SEQUENCE: SEQ ID NO:1
    33  Xaa Xbb Tyr Xcc Lys Thr Xbb Thr
    34   1                 5
E--> 36 <210> SEQ ID NO: SEQ ID NO:2
    37 <211> LENGTH: LENGTH: 8
E--> 38 <212> TYPE: TYPE: PRT
    39 <213> ORGANISM: ORGANISM:Synthetic
    41 <221> NAME/KEY: MOD_RES
    42 <222> LOCATION: (1)...(0)
    43 <223> OTHER INFORMATION: Xaa = D-Phe
E--> 44 <224> Xbb = Cys with an intramolecular disulfide bond
E--> 46 <225> Xcc = D-Trp
E--> 47 <226> Xdd = Thr-OH
E--> 49 <400> SEQUENCE: SEQ ID NO:2
    50  Xaa Xbb Tyr Xcc Lys Thr Xbb Xdd
    51   1                 5
E--> 52 <210> SEQ ID NO: SEQ ID NO:3
    53 <211> LENGTH: LENGTH: 11
E--> 54 <212> TYPE: TYPE: PRT
    55 <213> ORGANISM: ORGANISM:Synthetic
    57 <221> NAME/KEY: MOD_RES
```

see following pages for explanation

delete
~~SEQUENCE LISTING~~

delete
<110> ~~APPLICANT:~~ Samuel I. Achilefu
Raghavan Rajagopalan
Richard B. Dorshow
Joseph E. Bugaj
delete
~~ASSIGNEE:~~ Mallinckrodt Inc.

DO NOT insert alphabetical headings when using new Sequence Rules format. Just use numeric identifiers.

<120> ~~TITLE:~~ Versatile Hydrophilic Dyes

<130> ~~DOCKET/FILE REFERENCE:~~ MRD-67   → <140> <141> *insert here and response*

<150> ~~PRIOR APPLICATION NUMBER:~~ 09/484,321
<151> ~~FILING DATE:~~ 2000-01-18

<160> ~~NUMBER OF SEQUENCES:~~ 8

<170> ~~SOFTWARE:~~ FastSEQ for Windows Version 3.0

This numeric identifier is MANDATORY whenever <221>, <222>, or <223> is shown

<210> ~~SEQ ID NO:~~ 1   *Per new Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or Scientific name*
<211> ~~LENGTH:~~ 8
<212> ~~TYPE:~~ PRT
<213> ~~ORGANISM:~~ (Synthetic)   *see item 12 on Error Summary sheet (global error)*
<220>
<221> MOD_RES
<222> (1)...(0)   (1)...(7)   *use Xaa for all modified residues*
<223> Xaa = D-Phe
use <223> consistently   <224> (Xbb) = Cys with an intramolecular disulfide bond
between two Cys amino acids
<225> (Xcc) = D-Trp
Xaa

<400> ~~SEQ ID NO:~~ 1
Xaa (Xbb) Tyr (Xcc) Lys Thr (Xbb) Thr
1   Xaa   Xaa   5   Xaa

Since Xaa's at different locations represent different amino acids, state location in <223>, tov. e.g., "Xaa at locations 2 and 7 represent Cys with an intramolecular disulfide bond between two Cys amino acids."

<210> ~~SEQ ID NO:~~ 2
<211> ~~LENGTH:~~ 8
<212> ~~TYPE:~~ PRT
<213> ~~ORGANISM:~~ (Synthetic)   *item 12*
→ <220>
<221> MOD_RES
<222> (1)...(0)   (8)
<223> Xaa = D-Phe
<224> (Xbb) = Cys with an intramolecular disulfide bond
between two Cys amino acids
<225> (Xcc) = D-Trp
<226> Xdd = Thr-OH

<400> ~~SEQ ID NO:~~ 2
Xaa (Xbb) Tyr (Xcc) Lys Thr (Xbb Xdd)
1   5

same error as previous sequence

"Xaa at location 1 represents D-Phe."

```
<210> SEQ ID NO: 3
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12
<220>
<221> MOD_RES
<222> (1)...(0)

<400> SEQ ID NO: 3
Gly Ser Gly Gln Trp Ala Val Gly His Leu Met
1               5                   10


<210> SEQ ID NO: 4
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12
<220>
<221> MOD_RES
<222> (1)...(0)

<400> SEQ ID NO: 4
Gly Asp Gly Gln Trp Ala Val Gly His Leu Met
1               5                   10


<210> SEQ ID NO: 5
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12
<220>
<221> MOD_RES
<222> (1)...(0)

<400> SEQ ID NO: 5
Asp Tyr Met Gly Trp Met Asp Phe
1               5


<210> SEQ ID NO: 6
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12
<220>
<221> MOD_RES
<222> (1)...(0)  (6)

<400> SEQ ID NO: 6
Asp Tyr Nle Gly Trp Nle Asp Phe
1               5
```

same error

This is a modified residue; per Sequence Rules use Xaa in the sequence itself, and explain what it represents in <220>-<223> section. Do not use the modified residue in the sequence itself.

```
<210> SEQ ID NO: 7
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12

→ <220>
<221> MOD_RES
<222> (1)...(0)    (6)

<223> <228> Xff = D-Asp
                Xaa
<400> SEQ ID NO: 7
Xff Tyr Nle Gly Trp Nle Asp Phe
1               5            →  insert Xaa and explain

<210> SEQ ID NO: 8
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Synthetic    item 12
<220>          Xaa
<223> <229> Xgg = D-Lys

<400> SEQ ID NO: 8
Xgg Pro Arg Arg Pro Tyr Ile Leu
1               5
```

See sample Sequence Listing (attached) for valid format

Appendix A To Subpart C to Part I—Sample Sequence Listing

<110> Smith, John

Smith, Jane

<120> Example of a Sequence Listing

<130> 01-00001

<140> US 08/999,999

<141> 1998-02-28

<150> EP 91000000

<151> 1997-12-31

Please consult

<160> 2

<170> PatentIn ver. 2.0

<210> 1

<211> 403

<212> DNA

<213> Paramecium aurelia

<220>

<221> CDS

<222> 341..394

<300>

<301> Doe, Richard

<302> Isolation and Characterization of a Gene Encoding a

Protease from Paramecium sp.

<303> Journal of Fictional Genes

<304> 1

<305> 4

<306> 1 - 7

<307> 1988-06-20

<400> 1

```
ctactctact ctactctcat ctactatctt ctttggatct ctgagtctgc ctgagtggta   60

ctcttgagtc ctggagatct ctcctctcac atgtgatcgt cgagactgac cgatagatcg   120

ctgactgact ctgagatagt cgagcccgta cgagacccgt cgagggtgac agagagtggg   180

cgcgtgcgcg cagagcgccg cgccggtgcg cgcgcgagtg cgcggtgggc cgcgcgaggg   240

ctttcgcggc agcggcggcg ctttccggcg cgcgcccgtc cgcccctaga cctgagaggt   300

cttctcttcc ctcctcttca ctagagaggt ctatatatac atg gtt tca atg ttc   355
                                                Met Val Ser Met Phe
```

agc ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgtttgctc   403

Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val

                    10   15

<210> 2

<211> 18

<212> PRT

<213> Paramecium aurelia

<400> 2

Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu

1   5   10   15

Phe Val

ed: May 22, 1998.
· A. Lehman,
`ant Secretary of Commerce and
`issioner of Patents and Trademarks.
oc. 98-14194 Filed 5-29-98; 8:45 am]
3 CODE 3510-16-C

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	'Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials.	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOs	Count includes total number of SEQ ID NOs	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:#:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a seque...
<223>	Other Information	Other relevant information; four lines maximum	M, under the following Conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	O
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	O
<302>	Title		O
<303>	Journal		O
<304>	Volume		O
<305>	Issue		O
<306>	Pages		O
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	O
<308>	Database Accession Number	Accession number assigned by database including database name	O
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	O
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	O

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	O
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	O
<313>	Relevant Residues	FROM (position) TO (position)	O
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	M

5. Section 1.824 is revised to read as follows:

1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.821(e) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;

Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u> <u>SUGGESTED CORRECTION</u>   SERIAL NUMBER: _09/757,333_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped " down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering
The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length
Sequence(s) _____ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
sequence(s) _____. Normally, PatentIn would automatically generate this section from the
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
sections for Artificial or Unknown sequences.

8 _____ Skipped Sequences (OLD RULES)
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES)
Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 _____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES)
Sequence(s) _____ are missing this mandatory field or its response.

1-8

12 _____ Use of <220>Feature (NEW RULES)
Sequence(s) _____ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.